The Maddison Project

What Makes Maddison Right?

Maddison-Project Working Paper WP-7

Jan Luiten van Zanden and Debin Ma

September 2017

What Makes Maddison Right?

Jan Luiten van Zanden (University of Utrecht), Debin Ma (LSE)

Abstract

We discuss the recent paper by Deng and O'Brien 'Why Maddison was wrong' that criticizes the contribution by economic historians in general and Angus Maddison in particular who try to quantify economic performance of China in a way comparable with recent research on growth patterns in Europe (and other parts of the world). We argue that, whereas some estimates by Maddison may have been shown incorrect by later research, his basic intuition about China was not incorrect. Moreover, his research strategy, aimed at challenging colleagues to make better estimates, has been very productive and has produced valuable insights in the process of long term economic development.

Kent Deng and Patrick O'Brien recently published a review of the 'Great Divergence debate' in this journal, with the challenging title 'Why Maddison was wrong'. They presented their 'brief review' of this debate, which began in 2000 with the publication of Kenneth Pomeranz' (2000) seminal book 'The Great Divergence: Europe, China and the Making of the Modern World'. Their take on this debate was that Pomeranz maintained 'that standards of living afforded by the economy for populations contained within the political boundaries of the Ming-Qing Empire of China (1368-1911) did not fall behind het levels of well-being afforded to the populations of the national economies of Western Europe until late in the eighteenth century' (Deng and O'Brien 2017, 20-21). Or, the divergence between Western Europe and China only began in the final decades of the eighteenth century – after 1780. It is a rather restricted interpretation of this debate – which, others would argue, is also about the causes of the divergence that did happen, and about the relevant unit of comparison, but those parts of the debate are not covered by the review (and shall therefore not be considered here either). They then discuss the debate about this hypothesis in highly subjective terms: the 'novel theses' by Pomeranz were subject to 'bombardment' by 'the heavy artillery of statistical-cum-econometric 'tests''. It is this discussion that we take issue with.

During the past two decades, the Pomeranz book and related publications by his colleagues from the California School have led to a flowering of new research into the economic history of China, with the aim of testing these ideas. In at least four different ways the hypothesis has been subjected to more detailed quantitative research, and in all cases the conclusion has been that the Pomeranz hypothesis cannot be proven right and is probably incorrect. These four tests are: (1) the development of GDP and GDP per capita in China (and Europe), (2) the development of real wages (of unskilled labourers), (3) the evolution of agricultural productivity and the consumption level of the population, and (4) the rate of urbanization. All four literatures conclude that the Pomeranz hypothesis is probably too optimistic about the standard of living of Chinese population compared with that of the most advanced parts of Western Europe.

One of the paradoxes of the recent contribution by Deng and O'Brien is that in another paper they presented a detailed overview and analysis of the third test, concluding that 'our clarified and recalibrated estimates suggest that, from the early seventeenth century onwards, the state, institutions and foundational culture of the Chinese empire were failing to cope nearly as well with "the pressures of numbers and environmental degradation" as states and economies of Western Europe' (an interpretation which comes close to that of Peer Vries (2013) who is quite critical of Pomeranz). Or, as the authors write themselves, 'Our survey, critique, and recalibration of the data produced for this particular and altogether more promising line of historical enquiry rejects inference-derived statistical evidence that continues to be widely cited to support the revisionist claims of the California School. Recalibrated and tabulated here into kilocalories for purposes of reciprocal comparisons, that evidence suggests that although the population of early modern Jiangnan enjoyed "nutritional security," its standard of living was declining and falling to levels below standards afflicting England's laboring poor.' In sum, they are highly critical of the core hypotheses of the California School.

This is all valid economic historical research. But in the 'Why Maddison is Wrong' paper they develop an entirely different argument. The core argument is that 'the data for China accessible in secondary sources do not provide historical runs of estimates either for GDP or for total population, let alone for any purchasing-power-parity rates of exchange estimates' (p. 23). With this and similar arguments, they seek to criticize the conclusions from other research, focused on measuring GDP, real wages or the level of urbanization, that arrive at similar conclusions as their research on agricultural productivity and consumption levels. In particular Angus Maddison is the bad guy in this saga, as he produced estimates of GDP per capita which were based on limited evidence. Moreover, his approach – linking time series to a benchmark estimate for 1990 – has been criticized by 'a platoon of distinguished economists' who 'found his methods and estimates to be conceptually and statistically unacceptable as historical evidence' (Deng and O'Brien 2017, 21).

We firmly believe that critiques, assessments and summaries on the state of the Great Divergence debate and on the larger research agenda on assessing long-term living standards are very welcome and are also in the original spirit of the Maddison research. However, we feel that their characterization of the Maddison works and others research related to the issue are incomplete and inaccurate, and in some cases misleading. Firstly, the Maddison estimates are less about right or wrong than about trying to achieve better or best estimates by overcoming the current constrain on data and methodologies overtime. They are dynamic processes that require constant updating based on new data sources and methodologies. Secondly, the issues raised by Deng and O'Brien on the Maddison estimates by critiques have long been known to the scholars in the field of international comparison including Maddison himself and many of them are not peculiar to Chinese economic history. Indeed, some of their own summaries or assessments are not presented as meaningful or superior alternatives in term of either methodology or empirics, or in many ways not markedly different from the existing works by Maddison and others. [1] Hence, we are puzzled by the highly critical and sometimes dismissive tones on these works.

[1] For example, their use of calories in their recalibration is a useful cross-checks (Deng and O'Brian 2015). But the method of using calories is already part of the constructing consumption basket in the Allen et al 2011 paper, which they critiqued. The Allen et al 2011 consumption basket used a combination of a calories, proteins

Their paper starts out with their assessment on the use of nominal and real wages, which again draws from their critique of others' work mainly by Allen et al (Deng and O'Brien 2016).[2] The Allen et al 2011 concludes that real wages in Beijing, Canton and the Yantze delta were already in the 18th century significantly lower than in North Western Europe based on a large variety of sources from both the government (related for example to construction projects) and from companies (both foreign, like the VOC in Canton, and domestic firms). The Allen et al 2011 represents the first systematic attempt at constructing long-term wages for China and other Asian countries. So, we are happy to see this paper being given the kind of proper scrutiny from Deng and O'Brien (2016). As two co-authors on this paper, we find their meticulous critique very helpful to the debate despite numerous (perhaps unavoidable) misunderstandings. The team that worked on the Allen et al 2011 paper includes historians with a Chinese background and specialists on Chinese historical sources. The Allen et al 2011 team has taken great care to lay out their data sources and estimation procedures, which – we are happy to see – have allowed other scholars to trace, reproduce and critique the result. In fact, the major issues they discuss were already raised in the Allen et al 2011 paper; we alerted the readers to potential margins of errors, and in many cases, tested them with robustness checks and alternative scenarios.

For example, the main argument that is now used against the results of this research that labour markets in China were more marginal than in Western Europe. This may be correct (although there is perhaps a tension between this line of defense and the original argument of the Great Divergence book that institutional structures in China were very similar to those of Europe), but the Allen et al 2011 (p.29) paper compared the daily wage rates (of short-term laborers) with the bottom of the income pyramid and argues that the purchasing power of the wages of unskilled labourers may be good guide to the standard of living of a much larger group of lower incomes.

On similar grounds the new research into the development of the urbanization ratio, which shows a strong decline of this measure of economic transformation from the early Qing (c.12%) to the late 18th century (about half that level). So we have the results of four tests of the Pomeranz' hypothesis, based on different sets of sources, all with their own problems and limitations, but also all pointing into the same direction, that is, that this hypothesis (as summarized by Deng and O'Brien) is not correct. In fact, in their own original, data-based contribution, they arrive at the same conclusion.

Back to Maddison, and his contribution, where we have problems with the Deng and O'Brien characterization. Firstly, in the 1990s, way before the publication of Pomeranz book, Maddison began doing research on Chinese economic development, not because he wanted to address the 'Great Divergence' thesis, but because he was working on his grand synthesis of the growth of the world economy in the past two millennia, and China was obviously a very large part of that story. Very few people, however, were at the time (the1990s) doing quantitative economic historical research on China, which would result in the historical GDP series that he needed for

and actual consumption basket and commodities precisely because calories alone are an extremely limited indicator for measuring living standards.

[2] Strangely, the Allen et al 2011 paper, upon which much of their critique is based is either missing from their citation in this journal, or misquoted as Allen et al 2005, see p.22.

his synthesis. [3] Not shying away from the challenge, he decided to do the research himself. He travelled there, spoke to the most eminent economic historians, and tried to collect the data for his first set of estimates, making use of the state of the art literature. Maddison's OECD book on China contains a thorough re-evaluation of the entire 50 year Communist era GDP and their production accounts components, including separate PPP for agriculture and industry. For the pre-Communist era, Maddison examined a variety of works going way beyond Liu and Yeh 1965. Hence it is highly unfair to characterize "Madison utilized just two estimates for China's GDP in current prices:..".

For the earlier period, Maddison understandably relied on much scattered and dispersed evidences which include some population and agricultural output series all of which have their problems. In many cases, Maddison relied on his intellectual intuition based on a huge amount of reading on Chinese and global history sources to come up estimates for benchmark GDP series in 1990 prices. In numerous occasions, he explicitly termed them as "guess-estimates". These estimates were clearly not intended to be the final word in this regard. Maddison's research strategy was aimed at challenging people, and inviting them, if they knew better sources and methods, to make improved estimates of the development of GDP. He was very open to welcoming the result of all scholarly work in the field, with the aim to constantly improve the quality of the estimates of GDP, population and GDP per capita of his dataset. That this was a highly dynamic process, and that it was very successful in the long run, is clear from the dataset, which started, in the 1960s, with observations for a dozen or so industrial countries going back (sometimes) to 1870, and ended, in his synthesis of 14,000 data points, with estimates for more than a hundred countries covering, in a few cases, the entire two millennia. We agree that by publishing his 'guestimates' in tabulated form, Maddison made his data vulnerable to the risk of abuse by indiscriminant users who have not followed historical works carefully. But for people who did, the critiques by Deng and O'Brien on these estimates may be useful but rather superfluous.

The last decade has seen a new wave of works by various authors (XU et al 2016; Broadberry et.al. 2014) who produced papers with new estimates for the development of GDP during the Qing (and sometimes going back to Ming and Sung times). What all this new work has in common is that it is broadly consistent with the Maddison estimates made in his grand synthesis, and do not confirm the optimistic view of the California School. In fact, most recent estimates are more pessimistic and show a decline of GDP per capita during the period 1600-1800, where Maddison assumed stability. It seems in the end there is a lot of good things to be said about Maddison's initial intuition. Like Maddison's own works, all these works have various issues. The essence of the Maddison estimates was therefore that it was not supposed to be a static end-product of research, but that it had to be updated frequently in order to incorporate new research and improved estimates.

The second issue which Deng and O'Brien critiqued is the use of 1990 price benchmark. We should start out by saying that the issue of using 1990 benchmark for backward project is one issue that Maddison and his team have been keenly aware for a long-term and is part of larger

[3] Deng and O'Brien suggest that he published them in 2007, as a response to the Pomeranz book of 2000, he actually published his results in 1998, before the Great Divergence debate started (Maddison 1998).

index number problem for historical national accounts for all countries. The entire ICOP (International Comparison of Output and Productivity at: http://www.rug.nl/ggdc/projects/icop?lang=en), which was originally founded under Maddison's leadership at University of Groningen engaged in the systematic and large-scale construction of current price purchasing power parities (PPP) both as cross-checks as well as basis for new updates. Maddison himself has been highly encouraging of others conducting new current prices PPP parity estimates. Various research along this line actually reveal mixed results of current price versus backward extrapolated estimates. For example, Fukao, Ma and Yuan (2007) actually confirmed that the 1930s current price PPP estimates for China and Japan were not that far apart from the 1990 backward extrapolation used by Maddison. Li and Van Zanden (2012), for example, compared two most advanced part of Western Europe and China in the 1820s, and came to the conclusion that even the most urbanized parts of China clearly had a much lower level of GDP per capita than the Netherlands, a highly developed part of the European economy. So, in the end, Maddison stayed the use of a single 1990 benchmark backward extrapolation after weighing the costs and benefits of using single benchmark versus multiple benchmarks or current price estimates. Clearly, as we move further and further away from the 1990, it becomes increasingly necessary to make the updates. But it is matter of choice rather than right or wrong as Deng and O'Brian led us to believe.

When Maddison passed away in 2010 (it is a bit odd that Deng and O'Brien criticize him for not incorporating the new ICP results of 2011, which were published years after his death), it was decided that a group of scholars – close colleagues and friends - would set up 'the Maddison project' to continue his work (he had been very much in favor of this idea) (Bolt and Van Zanden 2014). This has resulted in a number of updates, the first one published in 2014, the second one now about to be launched, which have 'in his spirit' dealt with some of the problems mentioned by Deng and O'Brien (and probably by other members of the platoon as well). In the last update, the 1990 benchmark which has been criticized a lot, in combination with the method to extrapolate time series linked to this one benchmark back in time, has been replaced by a more sophisticated approach to combine benchmarks with time series, as developed by PennWorldTables (Feenstra, Inklaar and Timmer 2015). Both the 2014 and the 2017 updates include new estimates of GDP per capita for China, made by a team of authors (Xu, Shi, van Leeuwen, Ni, Zhang, and Ma (2016),also Mad and De Jong (20017)), which tend to arrive at results which are similar to the old set of 'guestimates' (it is again rather odd that Deng and O'Brien did not mention these works).

In the end, we want to argue that the research strategy which lay behind Maddisons figures is quite sound. As mentioned already, the years since 2000 have seen the emergence of the modern economic historical research that was carried out for other continents (Europe and North America in particular) since the rise of 'new economic history' in the 1960s and 1970s. Both Chinese and non-Chinese scholars have profited from the opening up of Chinese archives, the growth of universities in China, increased international academic exchange and the growing interest in the history of China, all resulting directly or indirectly from the spectacular economic performance of the country since about 1980. We are sure that this is just the beginning of a boom in the quantitative study of Chinese history. The Chinese state with its large bureaucracy has produced vast amounts of historical data – some of a quantitative nature – which scholars are now beginning to exploit. Of course, as Deng and O'Brien stress, all sources have their problems, and nothing can be taken at face value; but it is the core business of economic historians to deal with these problems, which are not unique for China (although Deng and

O’Brien want us to believe this is perhaps the case). Scholars working on European economic history have struggled with similar problems, and the many results of the quantification of European economic development are also still contested (as, for example, recent exchanges about economic growth in England demonstrate (Clark 2017)). European scholars have however been doing this kind of research for a much longer time, and have therefore built up more experience and credibility. Such a track record requires time and hard work, and it is perhaps natural that established scholars from older generations are critical of the new kid on the block – the ‘new economic history’ also met such resistance in the 1960s and 1970s (Drukker 2006).

Angus Maddison was never afraid being considered wrong especially when it comes to estimates for China, partly because he wanted to encourage or “provoke” other scholars with more specialist knowledge to come up with better estimates down the road. So, in that sense, he was right, perhaps not so much because he always estimated Chinese economic performance correctly (although his intuition seemed quite good in this regard as well), but because he developed the right research strategy to make consistent progress in the extension and refinement of his historical dataset of GDP and GDP per capita. This is his true achievement.

References

Allen, Robert C., Jean-Pascal Bassino, Debin Ma, Christine Moll-Murata, and Jan Luiten van Zanden (2011).‘Wages, prices, and living standards in China, 1738-1925: in comparison with Europe, Japan, and India’, *Economic History Review* 64, pp. 8-38.

Bolt, J. and J.L. van Zanden (2014). “The Maddison Project: collaborative research on historical national accounts”, *Economic History Review* 67, pp. 627-651.

Broadberry, S., Guan, H.H. and Li, D. D.(2014) China, Europe and the great divergence: a study in historical national accounting. [Accessed 6 Jan 2016.] Available from URL:http://eh.net/eha/wp-content/uploads/2014/05/

Clark, G. (2017), Growth or stagnation? Farming in England, 1200–1800. The Economic History Review. doi:10.1111/ehr.12528

Deng, Kent and O'Brien, Patrick (2015) Clarifying data for reciprocal comparisons of nutritional standards of living in England and the Yangtze Delta (Jiangnan), c.1644 – c.1840.Journal of World History, 26 (2) 233-267.

Deng, Kent and O’Brien, Patrick (2016). “Establishing Statistical Foundations of a Chronology for the Great Divergence: a survey and critique of the primary sources for the construction of relative wage levels for Ming-Qing China. *Economic History Review* 69, 4. Pp.1057-82.

Deng, Kent and Patrick O’Brien (2017), Why Maddison was wrong. The Great Divergence between Imperial China and the West. World Economics, 2 (18) 21-41.

Drukker, J.W. (2006) A Revolution that Bit its Own Tail. Amsterdam University Press.

Feenstra, R.C., R. Inklaar and M.P. Timmer, (2015), “The Next Generation of the Penn World Table” American Economic Review 105(10).

Fukao., Kyoji, Debin Ma and Tangjun Yuan “Real GDP in Pre-War East Asia: a 1934-36 Benchmark Purchasing Power Parity Comparison with the U.S.” (co-authored with), *Review of Income and Wealth,* Vol. 53 Issue 3, Sept. 2007, p503-537.

Li, B., and Van Zanden, J.L. (2012). Before the Great Divergence? Comparing the Yangzi Delta and the Netherlands at the Beginning of the Nineteenth Century. The Journal of Economic History, 72 (4), 956-989.

Ma, Ye, and H.J. de Jong (2016). Unfolding the turbulent century; a reconstruction of Chinas Historical National Accounts, 1840–1912. Review of Income and Wealth, 2017.

Maddison, Angus (1998) Chinese Economic Performance in the Long-Run. OECD Development Centre, Paris, 1998.

Pomeranz, Kenneth (2000). The Great Divergence. China, Europe and the Making of the Modern World Economy. Princeton: Princeton U. P.

Xu, Y. Z. Shi, B. van Leeuwen, Y Ni, Z Zhang, and Y Ma, (2016) 'Chinese National Income, ca. 1661-1933', *Australian Economic History Review* (forthcoming)

Xu, Yi, B. van Leeuwen and J.L. van Zanden, ‘Urbanization in China, ca. 1100–1900’. CGEH working paper 63.

Vries, Peer (2013) Escaping Poverty. The Origins of Modern Economic Growth. Vienna.